Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 2796–2800, May 1984
Developmental Biology

Molecular cloning of region-specific chorion-encoding RNA sequences

(cell differentiation/aeropyle crown/silkmoth)

JEROME C. REGIER, ANTONIS K. HATZOPOULOS, AND ANN CATHRALL DUROT

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60201

Communicated by E. Margoliash, December 27, 1983

ABSTRACT **We have constructed a cDNA clone library from poly(A)$^+$ RNA of very-late-period choriogenic silkmoth follicles. Clone DNAs that hybridize preferentially to RNA from the aeropyle crown region of the follicle (versus the flat region) were selected, and all could be placed in one of two homology groups. The two groups represent sequences encoding the very-late-period chorion proteins E1 and E2; this was established by hybrid-selected translation coupled with specific antibody precipitation. Regionalized synthesis of chorion proteins is restricted to the very late period, and its control can now be studied at the nucleic acid level.**

Choriogenesis in silkmoths follows a highly regulated temporal program of gene product expression (1, 2). Five developmental periods—very early, early, middle, late, and very late—are recognized by peak synthesis of unique subsets of proteins and by special modes of morphogenesis (3–5). Control of this temporal program has been analyzed in some detail previously, and it appears that changing temporal patterns of chorion protein synthesis are driven by corresponding changes in concentrations of encoding RNAs (6, 7). Until the end of the late period, most of these proteins are produced in all regions of the follicular epithelium, presumably by all cells. However, during the very late period, when the final 10% of chorion proteins are deposited, dramatic biochemical and morphological differences between two major regions of the follicular epithelium and associated chorion become apparent (8, 9). In the aeropyle crown region, which forms a band circumscribing the follicle, a novel subset of chorion proteins greatly accelerates its synthesis while synthesis of late-period proteins tails off. These very-late-period proteins assemble extracellularly into protruding respiratory structures of the chorion called aeropyle crowns. Two substructures of aeropyle crowns can be recognized with the electron microscope—lamellar chorion, which forms the protruding prongs of the crown and is similar in ultrastructure to the underlying chorion found in all regions, and filler, which is found within the cavity of the crown and underlying aeropyle channel and is very distinct in ultrastructure from lamellar chorion. Selected very-late-period chorion proteins that assemble specifically to form filler have been identified previously and are called E1 and E2 (9). Presumably, other very-late-period proteins of the A-, B-, and C-size classes assemble to form lamellar chorion. Flanking the aeropyle crown region are two flat sides called the flat region. In this region decreasing late-period protein synthesis is not accompanied by a corresponding dramatic increase in very-late-period protein synthesis. Consequently, no aeropyle crowns form, although small amounts of filler are synthesized and accumulate within underlying aeropyle channels.

Molecular cloning of very-late-period-specific nucleic acid sequences would permit further analysis of the molecular basis of regionalization, the organization of these particular cell-specific sequences in the chromosome, and the structural and evolutionary features of the encoded protein sequences. In this report, we describe the construction of a very-late-period cDNA library and the selection and partial characterization of two clones that contain encoding sequences of known aeropyle crown components.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Follicle Preparation. Choriogenic follicles from developing adult *Antheraea polyphemus* were isolated and staged as described (1, 2). To isolate pure aeropyle crown and flat region cells and underlying chorion, freshly dissected follicles were placed in 95% ethanol and cut in half. The oocyte and vitelline membrane were removed. Cellular regions were isolated in 95% ethanol under a dissecting microscope using a fine scalpel.

RNA Preparation. Total cellular RNA was isolated from follicle cells as described (6), with the modification that very-late-period follicles were mixed in a Vortex in lysis buffer rather than homogenized to avoid rupturing the oocyte. RNA used for molecular cloning was purified by two rounds of oligo(dT)-cellulose chromatography. RNA that was end labeled with [γ-^{32}P]ATP was first fragmented by incubation with 0.1 M Tris·HCl (pH 9.5 at 23°C) for 12 min at 80°C.

Protein Analysis. Follicles were labeled in tissue culture with [^{3}H]leucine and their protein products were analyzed on NaDodSO$_4$ gels as described (10, 11). E1- and E2-specific antibodies in rabbit serum were the gift of Ruth Griffin Shea. E1 and E2 antigens were isolated as described (8). Antibody precipitations were carried out in the presence of 0.1% NaDodSO$_4$ to minimize nonspecific precipitation.

Construction of the cDNA Library. Only the oldest follicle (approximately stage Xd) from each ovariole of multiple animals was used as a source for RNA in the construction of the library. The protocols used for synthesizing double-stranded cDNA, for hybridizing it to *Pst* I-linearized pBR322 DNA and for transforming *Escherichia coli* strain RRI are described elsewhere (12). Transformants were picked individually and transferred to microtiter wells.

Hybridization Analysis. Techniques for colony hybridization (13, 14), for dot hybridization (15), for RNA blot hybridization (16) using formaldehyde gels (17), and for hybrid-selected translation (18) are described elsewhere. The sensitivities of the hybrid-selected translations and dot hybridizations were substantially improved by binding plasmid inserts rather than whole linearized plasmids to nitrocellulose filters. Inserts were isolated from agarose gels (see method 2 in ref. 19) and purified through a cellulose acetate filter and an Elutip-d column (Schleicher & Schuell) prior to hybrid-selection or nick-translation. Cell-free translation mixtures used were commercially available wheat germ and reticulocyte lysates (Bethesda Research Laboratories).

RESULTS

Protein Synthesis During the Very Late Period. The very late period extends over the last quarter of choriogenesis (approximately 12 hr), during which time there are continuous changes in the pattern of protein synthesis, reflecting a shift from synthesis of predominantly late-period (stages VI–IX) to very-late-period (stages Xa–Xd) proteins. During stages Xa and Xb, synthesis of very-late-period-specific proteins increases but still accounts for a small fraction of total follicle protein synthesis (Fig. 1, lanes 1 and 2). By stage Xc, the majority of synthesis is very-late-period-specific (Fig. 1, lanes 3 and 4). By stage Xd, two very-late-period-specific proteins—E1 and E2—account for most follicle synthesis (Fig. 1, lanes 5 and 5′).

Construction of a cDNA Library and Selection of Region-Specific Clones. A cDNA clone library was constructed using poly(A)$^+$ RNA isolated from whole very-late-stage chorionating follicles. From the library of 3042 clones, 1128 clones were screened for preferential hybridization to RNA from the aeropyle crown region relative to that from the flat region, on the assumption that regionalization of very-late-period sequences might be found at the RNA level as well as at the protein level. A representative result is shown in Fig. 2 (*Left*), where end-labeled aeropyle crown region RNA isolated from very-late-stage follicles was hybridized to 95 lysed bacterial clone colonies. Approximately 40% of these showed various degrees of hybridization above background, indicating the presence of homotypic sequences in the RNA probe. A replica of the same bacterial colonies was also probed with end-labeled flat region RNA (Fig. 2 *Right*) isolated from the same very-late-stage follicles. Again, selected clones hybridize above background levels, and these can be placed into one of two categories. The much larger group, of which 12 are indicated by filled arrows, hybridize more intensely to aeropyle crown region RNA than to flat region RNA. The other group, three of which are indicated by open arrows, hybridize more intensely to flat region RNA.

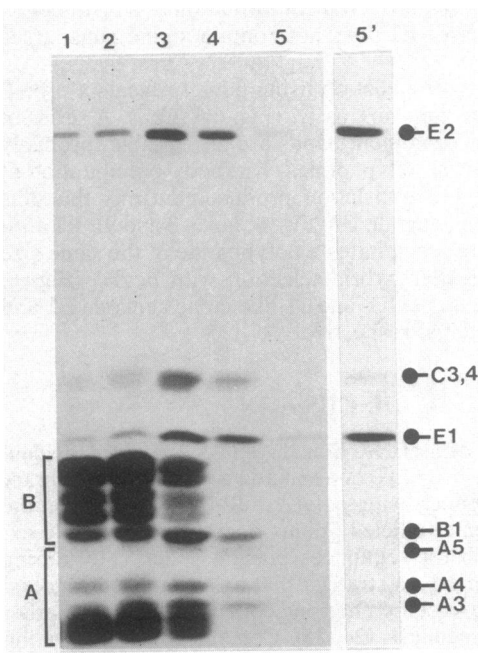

FIG. 1. *In vivo* patterns of protein synthesis during the very late period. Very-late-period follicles were labeled with [^{3}H]leucine, and the polypeptides were fractionated on NaDodSO$_4$ gels and visualized by fluorography. Equivalent amounts of protein were layered on all lanes. Lanes 1–5 were exposed to film for the same length of time; lane 5′ was exposed 10 times longer. A- and B-size proteins are identified on the left and very-late-period-specific proteins are identified on the right. Lanes: 1, stage Xa; 2, stage Xb; 3, stage Xc; 4, Xc$^+$; 5 and 5′, stage Xd.

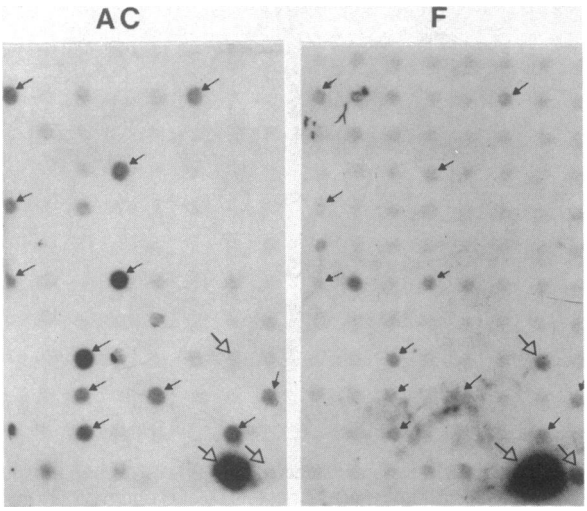

FIG. 2. Screening of the cDNA library with end-labeled very-late-period RNA isolated from different follicle regions. Two filter replicas containing 95 bacterial clones each (including plasmid vector in the lower lefthand corner) were hybridized to equal amounts of end-labeled aeropyle crown (AC) or flat (F) region RNA isolated from approximately stage Xa follicles. Hybridization conditions include 50% formamide, 0.45 M NaCl, probe at 2×10^6 cpm/ml, 42°C overnight. Solid arrows (both filters) identify selected clones that hybridize preferentially to aeropyle crown region RNA. Open arrows identify clones that hybridize preferentially to flat region RNA.

The first group of clones are reasonable candidates for encoding aeropyle crown region-specific proteins. A subset of these clones was placed in groups of sequence relatedness, based on the ability of selected plasmid inserts to cross-hybridize with other clones. Of four clone inserts tested, all hybridized with multiple clones, some of which were the same for different inserts. The overlap between cross-hybridization groups was increased by selecting clones with long inserts. Ultimately, two distinct groups of clones were defined on the basis of their cross-hybridization to all of 36 randomly chosen clones that encode abundant and region-specific RNAs, including the 12 identified by filled arrows in Fig. 2. Of the 36 clones shown in Fig. 3, 28 hybridized intensely to a representative insert (called pcvl 11) from one group and 8, to that from the other (called pcvl 3). There was no evidence for cross-hybridization between these two groups of clones, even at a relatively low criterion. Thus, it would appear that these two sequences represent distinct abundant RNA types present in very late stage follicles.

Clones with long inserts of each sequence type were selected. Inserts from pcvl 3 and three other pcvl 3-like clones were the longest (760 base pairs) out of 30 tested. In the other group, the pcvl 11 insert was the longest (2080 base pairs) from 31 clones tested.

Further Characterization and Identification of Region-Specific Clones. The colony hybridization results that indicated regional specificity (see Fig. 2) were confirmed on these two sequence types using the more sensitive technique of dot hybridization (Fig. 4). In the particular experiment, quantitation by scintillation counting showed that equivalent amounts of pcvl 3 and pcvl 4 (a pcvl 11 homolog) DNAs hybridized with 37 and 12 times as much aeropyle crown region RNA as flat region RNA, respectively.

RNA blot analysis using the two very-late-period sequence types as probes demonstrated hybridization to individual very-late-period RNA bands (Fig. 5). pcvl 1 (a pcvl 3 homolog) hybridized to a band significantly larger than

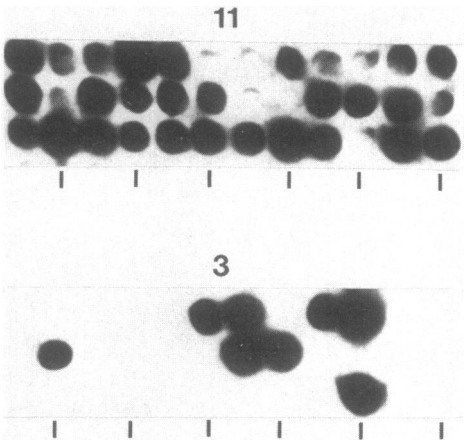

FIG. 3. Hybridization of selected aeropyle crown region-specific cDNA clones with inserts from pcvl 3 (*Lower*) and pcvl 11 (*Upper*). Replicate filters of 36 aeropyle crown region-specific cDNA clones (three rows of 12 dots in each panel) and the plasmid vector (no detectable hybridization; dot not shown) were probed with inserts from region-specific clones. Hybridization conditions included 0.45 M NaCl, poly(G) and poly(C) at 100 μg (each)/ml, fragmented pBR322 DNA at 0.4 μg/ml, nick-translated probes at 5×10^5 cpm/ml, 65°C overnight.

pc401 RNA (approximately 850 nucleotides versus 604 for pc401 RNA; see ref. 20); the latter encodes the largest known class B protein (21). pcvl 4 hybridized to a band of approximately 2000 nucleotides. From this RNA blot analysis, it would appear that both pcvl 3 and pcvl 11 contain near full-length copies of the corresponding RNAs and that neither cross-hybridizes to a significant degree with each other or with late-period-specific A and B family sequences, which were present in significant concentrations in the blotted RNA (unpublished observations).

On the basis of their regional specificity and their sizes (which exceed those expected for A- and B-family sequences), we reasoned that the pcvl 3 and 11 inserts might represent class E sequences. We have tested this hypothesis using positive hybrid-selection coupled with cell-free translation and specific antibody precipitation (Fig. 6). The specificities of the E2 and E1 antibodies were established by dem-

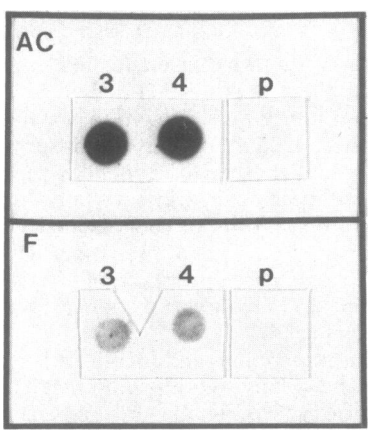

FIG. 4. Dot hybridization of purified aeropyle crown region-specific clone DNAs to end-labeled very-late-period RNA. Filter-bound linearized plasmid DNAs isolated from two aeropyle crown region-specific cDNA clones (3, pcvl 3; 4, pcvl 4) were hybridized to equal amounts of end-labeled regional RNAs (AC, aeropyle crown region RNA; F, flat region RNA), along with plasmid vector DNA (p) as control. RNA was isolated from stage Xa follicles. Hybridization conditions included 50% formamide, 0.45 M NaCl, fragmented pBR322 DNA at 2 μg/ml, probes at 7.5×10^7 cpm/ml, 42°C for 45 hr.

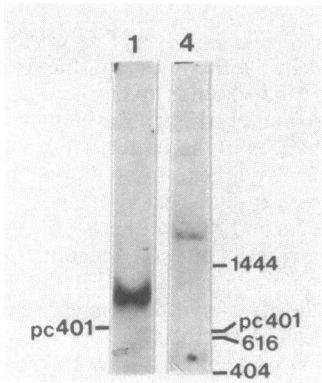

FIG. 5. RNA blot analysis of very-late-period RNA with aeropyle crown region-specific clone DNAs. Size-fractionated very-late-period poly(A)$^+$ RNA was transferred to nitrocellulose and probed with linearized region-specific clone DNAs (1, pcvl 1; 4, pcvl 4). Filters were also probed with a known class B clone DNA (pc 401). Single-stranded end-labeled DNA markers were run in parallel and their units of length are given in nucleotides (1444, 616, 404). Hybridization conditions included 50% formamide, 0.45 M NaCl, poly(A) at 10–40 μg/ml, probes at 5×10^6 cpm/ml (lane 1) and 4.6×10^5 cpm/ml (lane 4), 42°C overnight.

onstrating that only single bands of the correct size precipitate from a mixture of follicle proteins synthesized *in vivo* in very-late-period follicles (Fig. 6: E2, lanes 1–3; E1, lanes 1′–3′). Likewise, when very-late-period RNA is translated in a cell-free system, only specific translation products of similar sizes are immunoprecipitated (Fig. 6: E2, lanes 4–6; E1, lanes 4′–6′). The E1 cell-free translation product appears slightly heterogeneous in size, consistent with either population polymorphism or gene multiplicity (22, 23). It is interesting that the E1 cell-free translation product does not appear to be larger than its *in vivo* counterpart (unlike A and B family chorion polypeptides; ref. 9), perhaps reflecting an unusual signal peptide or overall conformation or a distinct glycosylation pattern. (E1 does not contain amino sugars; see ref. 8.)

Hybrid-selected translation using pcvl 4 reveals a prominent band of the same size as E2 (Fig. 6, lane 7). A series of smaller and less prominent bands are also visible, probably representing partial polypeptides. Antibody precipitation of the pcvl 4-selected translation product confirms that this clone does indeed encode E2 (Fig. 6; lanes 8 and 9). E1 antibody specifically precipitates a polypeptide of the same size as is translated after hybrid-selection with pcvl 3 (Fig. 6, lanes 7′–9′). Thus, pcvl 3- and 11-like clones encode E1 and E2 chorion proteins, respectively.

DISCUSSION

Construction of a cDNA Library and Selection of Region-Specific Clones. We have constructed a cDNA clone library of follicle sequences using poly(A)$^+$ RNA isolated from the oldest choriogenic follicles (minus oocyte) present. We expected to enrich for sequences present specifically during the very late period—a strategy that has been used successfully for cloning specific chorion sequences present at other periods of choriogenesis (24, 25). Conclusive demonstration that there were only two cross-hybridizing groups represented in the most abundant, regionalized RNA sequences was not immediately obvious, since it turned out that a number of E2-encoding clones did not cross-hybridize. Only when a near full-length clone copy had been identified was it possible to demonstrate that nonoverlapping portions of the same E2 sequence were being cloned separately. This was not a problem for E1 clones, presumably because of the substantially shorter E1 RNA sequence.

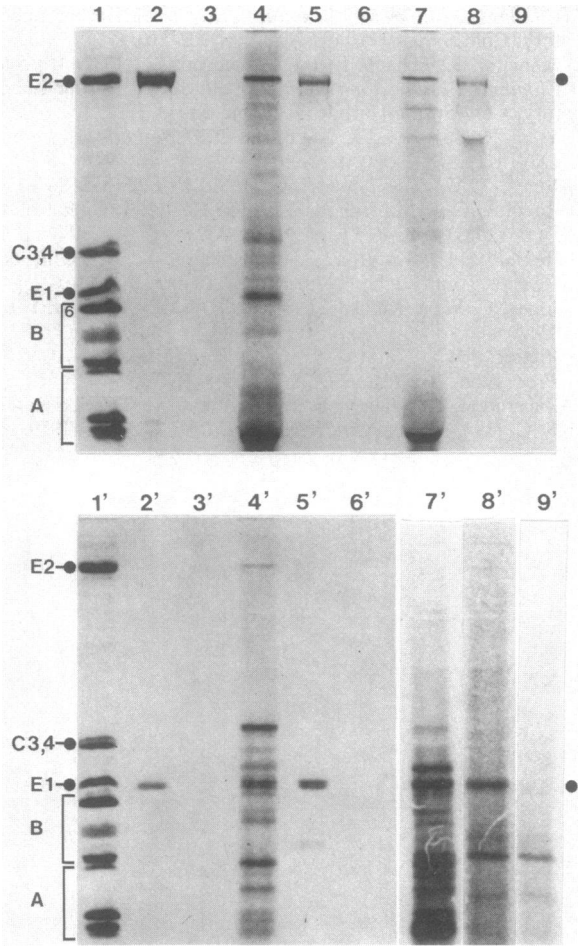

FIG. 6. Identification of region-specific cDNA clone homology groups as E1 and E2 encoding. *In vivo*-synthesized stage Xc-specific follicular proteins (lanes 1 and 1′) were immunoprecipitated with E2 (lane 2)- or E1 (lane 2′)-specific antibody or with their pre-immune serum (lanes 3 and 3′, respectively) and then fractionated on Na-DodSO$_4$ gels and fluorographed. Very-late-period RNA (lanes 4–6, total RNA from stage Xd follicles; lanes 4′–6′, total aeropyle crown region RNA from stage Xa follicles) was translated and then fractionated, either directly (lanes 4 and 4′) or after immunoprecipitation with E2 (lane 5)- or E1 (lane 5′)-specific antibody or with pre-immune serum (lanes 6 and 6′, respectively). Stage Xd RNA was hybrid-selected with pcvl 4 DNA, translated, and fractionated, either directly (lane 7) or after immunoprecipitation with E2-specific antibody (lane 8) or pre-immune serum (lane 9). Poly(A)$^+$ RNA from approximately stage Xa follicles was either translated directly (lane 7′) or else was hybrid-selected with pcvl 3 DNA and then translated (lane 8′). Endogenous synthesis is shown in lane 9′. Wheat germ and reticulocyte lysate-based translation systems were used for the identification of E1 and E2, respectively. The faint bands in lanes 8 and 5′ have not been analyzed further, although they could represent premature termination products of E2 and E1, respectively. Chorion protein classes and selected subclasses are identified on the left. Filled circles on the right identify E2 and E1 proteins.

E1 and E2 Sequences Are Abundant and Regionalized During the Very Late Period. Synthesis of aeropyle crown region-specific proteins becomes faintly detectable during the middle period but major only during the very late period (Fig. 1; see also ref. 9). Even during the first half of the very late period (stages Xa and Xb), region-specific synthesis is relatively minor. This changes dramatically near stage Xc and, by stage Xd, E1 and E2 proteins account for almost all synthesis.

The choriogenic follicles from which RNA was extracted for subsequent cDNA cloning were developmentally the oldest present—at or near stage Xd. Thus, it is not surprising that all of the most abundant clones selected for further analysis were E1 or E2 encoding. This parsimony between specific RNA abundance and level of specific protein synthesis is completely consistent with earlier findings for early, middle, and late-period-specific sequences (6, 7).

Regionalization of pcvl 3- and pcvl 11-like sequences was first demonstrated by colony hybridization (Fig. 2) and then confirmed by dot hybridization (Fig. 4). Both sequence types are more than 10 times as abundant in the aeropyle crown region as in the flat region. On the basis of this finding, it seems likely that differing regional rates of very-late-period protein synthesis are driven by corresponding regional differences in corresponding mRNA concentrations, although a more definitive study at multiple stages of development, using earlier period-specific chorion sequences as controls is clearly required. E1 and E2 sequences are present above background levels in the flat region as well as in the aeropyle crown region, consistent with the previous observation of a small amount of very-late-period-specific protein synthesis in the flat region, including E1 and E2.

An unexpected finding was that a few clones hybridized more extensively to flat region RNA than to aeropyle crown region RNA (Fig. 2), despite the fact that no chorion proteins are unique to the flat region, nor even relatively enriched (8). Tentatively, we suggest that late period RNA sequences originally present in both cellular regions persist somewhat longer into the very late period in flat region cells than in aeropyle crown region cells.

E1 and E2 Sequences Are Distinct from Each Other and from Previously Characterized Chorion Sequences. Four lines of evidence suggest that E1 and E2 sequences are distinct from each other and from previously characterized late-period-specific class A and B family sequences. (*i*) Near full-length inserts from pcvl 3 and pcvl 11 hybridize at low stringency to distinct sets of cDNA clones, showing no cross-hybridization, even when the autoradiograms are overexposed (Fig. 3). (*ii*) pcvl 3- and pcvl 11-like inserts hybridize to distinct and individual RNA bands (Fig. 5). No other hybridizing RNA bands are detectable. (*iii*) The E1- and particularly E2-encoding RNA bands are larger than any previously characterized A- or B-specific RNA. (*iv*) Finally, E1- and E2-specific antibody precipitations of both *in vivo*-synthesized proteins and cell-free translated proteins (with or without prior hybrid-selection of encoding mRNA) show that only individual protein bands of class E sizes are immunoprecipitated (Fig. 6). More extensive comparative studies, including sequence analysis, will be needed to determine whether lesser degrees of sequence homology with each other and with other chorion sequences occur.

We thank Ruth Griffin Shea for giving us specific antibodies to E1 and E2 chorion proteins and F. C. Kafatos for comments on the manuscript. This work was supported by grants from the National Institutes of Health and the American Cancer Society, Illinois Division.

1. Paul, M., Goldsmith, M. R., Hunsley, J. R. & Kafatos, F. C. (1972) *J. Cell Biol.* **55**, 653–680.
2. Paul, M. & Kafatos, F. C. (1975) *Dev. Biol.* **42**, 141–159.
3. Regier, J. C., Mazur, G. D., Kafatos, F. C. & Paul, M. (1982) *Dev. Biol.* **92**, 159–174.
4. Mazur, G. D., Regier, J. C. & Kafatos, F. C. (1982) in *Insect Ultrastructure*, eds. Akai, H. & King, R. C. (Plenum, New York), pp. 150–185.
5. Regier, J. C. & Kafatos, F. C. (1984) in *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, eds. Kerkut, G. A. & Gilbert, L. I. (Pergamon, Oxford), Vol. 1, in press.
6. Thireos, G. & Kafatos, F. C. (1980) *Dev. Biol.* **78**, 36–46.
7. Sim, G. K., Kafatos, F. C., Jones, C. W., Koehler, M. D., Efstratiadis, A. & Maniatis, T. (1979) *Cell* **18**, 1303–1316.
8. Regier, J. C., Mazur, G. D. & Kafatos, F. C. (1980) *Dev. Biol.* **76**, 286–304.

9. Mazur, G. D., Regier, J. C. & Kafatos, F. C. (1980) *Dev. Biol.* **76,** 305–321.
10. Goldsmith, M. R., Rattner, E. C., Koehler, M. D., Balikov, S. R. & Bock, S. C. (1979) *Anal. Biochem.* **99,** 33–40.
11. Lasky, R. A. & Mills, A. D. (1975) *Eur. J. Biochem.* **56,** 335–341.
12. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 230–242, 254, 255.
13. Grunstein, M. & Wallis, J. (1979) *Methods Enzymol.* **68,** 379–389.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), p. 314.
15. Kafatos, F. C., Jones, C. W. & Efstratiadis, A. (1979) *Nucleic Acids Res.* **7,** 1541–1552.
16. Thomas, P. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
17. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 202–203.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 332–333.
19. Yang, R. C.-A., Lis, J. & Wu, R. (1979) *Methods Enzymol.* **68,** 178–181.
20. Jones, C. W. & Kafatos, F. C. (1980) *Cell* **22,** 855–867.
21. Rodakis, G. C., Moschonas, N. K., Regier, J. C. & Kafatos, F. C. (1983) *J. Mol. Evol.* **19,** 322–332.
22. Regier, J. C., Kafatos, F. C., Kramer, K. J., Heinrikson, R. L. & Keim, P. S. (1978) *J. Biol. Chem.* **253,** 1305–1314.
23. Jones, C. W. & Kafatos, F. C. (1981) *Mol. Cell. Biol.* **1,** 814–828.
24. Regier, J. C., Kafatos, F. C. & Hamodrakas, S. J. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1043–1047.
25. Lecanidou, R., Eickbush, T. H., Rodakis, G. C. & Kafatos, F. C. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1955–1959.